



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2018

Crop classification in a heterogeneous arable landscape using uncalibrated UAV data

Böhler, Jonas E ; Schaepman, Michael E ; Kneubühler, Mathias

Abstract: Land cover maps are indispensable for decision making, monitoring, and management in agricultural areas, but they are often only available after harvesting. To obtain a timely crop map of a small-scale arable landscape in the Swiss Plateau, we acquired uncalibrated, very high-resolution data, with a spatial resolution of 0.05 m and four spectral bands, using a consumer-grade camera on an unmanned aerial vehicle (UAV) in June 2015. We resampled the data to different spatial and spectral resolutions, and evaluated the method using textural features (first order statistics and mathematical morphology), a random forest classifier for best performance, as well as number and size of the structuring elements. Our main findings suggest the overall best performing data consisting of a spatial resolution of 0.5 m, three spectral bands (RGB—red, green, and blue), and five different sizes of the structuring elements. The overall accuracy (OA) for the full set of crop classes based on a pixel-based classification is 66.7%. In case of a merged set of crops, the OA increases by 7% (74.0%). For an object-based classification based on individual field parcels, the OA increases by 20% (OA of 86.3% for the full set of crop classes, and 94.6% for the merged set, respectively). We conclude the use of UAV to be most relevant at 0.5 m spatial resolution in heterogeneous arable landscapes when used for crop classification.

DOI: <https://doi.org/10.3390/rs10081282>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-156769>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Böhler, Jonas E; Schaepman, Michael E; Kneubühler, Mathias (2018). Crop classification in a heterogeneous arable landscape using uncalibrated UAV data. *Remote Sensing*, 10(8):1282.

DOI: <https://doi.org/10.3390/rs10081282>

Article

Crop Classification in a Heterogeneous Arable Landscape Using Uncalibrated UAV Data

Jonas E. Böhler *, Michael E. Schaepman  and Mathias Kneubühler 

Department of Geography, Remote Sensing Laboratories (RSL), University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland; michael.schaepman@geo.uzh.ch (M.E.S.); mathias.kneubuehler@geo.uzh.ch (M.K.)

* Correspondence: jonas.boehler@geo.uzh.ch; Tel.: +41-446-355-249

Received: 22 June 2018; Accepted: 10 August 2018; Published: 14 August 2018



Abstract: Land cover maps are indispensable for decision making, monitoring, and management in agricultural areas, but they are often only available after harvesting. To obtain a timely crop map of a small-scale arable landscape in the Swiss Plateau, we acquired uncalibrated, very high-resolution data, with a spatial resolution of 0.05 m and four spectral bands, using a consumer-grade camera on an unmanned aerial vehicle (UAV) in June 2015. We resampled the data to different spatial and spectral resolutions, and evaluated the method using textural features (first order statistics and mathematical morphology), a random forest classifier for best performance, as well as number and size of the structuring elements. Our main findings suggest the overall best performing data consisting of a spatial resolution of 0.5 m, three spectral bands (RGB—red, green, and blue), and five different sizes of the structuring elements. The overall accuracy (OA) for the full set of crop classes based on a pixel-based classification is 66.7%. In case of a merged set of crops, the OA increases by ~7% (74.0%). For an object-based classification based on individual field parcels, the OA increases by ~20% (OA of 86.3% for the full set of crop classes, and 94.6% for the merged set, respectively). We conclude the use of UAV to be most relevant at 0.5 m spatial resolution in heterogeneous arable landscapes when used for crop classification.

Keywords: consumer-grade camera; land cover; very high resolution (VHR); random forest (RF) classifier; object-based classification; pixel-based classification; small-scaled agricultural fields; texture; spatial feature

1. Introduction

Agriculture reacts very sensitively to climate change [1]. Since the world population is expected to grow to 9.6 billion by the year 2050 [2], global food demand is increasing [3], and therefore, the availability of accurate and timely information about agriculture on a global [4], as well as on a local scale, is essential [5] to ensure that a growing world population can be fed. In order to address the problems of food insecurity or the volatility of the food market, remote sensing technologies provide a wide range of opportunities to assess these challenges [4,6].

Numerous aspects in agriculture and agronomy have been addressed for years with the support of remote sensing [6]. Among others are estimation of yield [7], biomass [8], seasonal plant development and stress monitoring [9,10], phenology and vigor [11], and land cover or land use change [12,13].

Accurate land cover assessments form the basis for such analyses in agricultural areas [14], and are particularly important for planning of water resources [15], automated short-term monitoring for yield estimation [16], sustainable land management [17], crop modeling before the end of season [18], or plot extraction for high-throughput phenotyping [19–21]. Further, current conditions and extent of land cover are needed as a basis for climate change modeling [22]. Often, information about arable land

cover is only available in administrative units, while detailed information on spatial distribution is lacking [23] or not accessible to the general public for reasons of data protection regulations [24].

In Switzerland, the agricultural landscape is mainly characterized by small structured farmland [25]. Support of farmers through direct payments plays a key role in Swiss agricultural policy. In 1996, the policy was changed from underwriting guarantee for most of the agricultural products towards compensation for provided services [26]. Today, the focus of direct payments to farmers in Switzerland is aligned with preservation of the cultural landscape, ensuring food supplies, sustainable production, enrichment of biodiversity, and contributions to the quality of the landscape. Small-scale subdivision of the arable area into portions of different crops guarantees an attractive landscape and fosters biodiversity [26,27]. Finally, direct payments are based on acreage, compensation for hilly terrain, individual crop plants, and their respective combinations [28]. Contributions for specific crops are paid, among others, for oilseed and sugar beet, while different types of cereals are not of specific interest.

The respective reporting is so far based on self-declaration, and likewise, the controlling is based on field visits. Accurate verification of crop types is important for subsidy control, to monitor biodiversity relevant land management practices, like crop rotation, soil protection and sustainable application of fertilizer, and to link up protected areas and ecological compensation areas [26].

Sharing information from reporting and controlling is often time consuming, and not straightforward. This is critical when it comes to the need for early warning [11], e.g., in case of water shortage, when sustainable planning is necessary [15]. Future predictions by climate models indicate that water will become a limited resource in Switzerland, as well [22,29]. Since water need for irrigation is crop dependent, accurate and timely information on crop allocation and acreage is required for planning purposes.

With the agricultural landscape in Switzerland being small structured, data of spatially very high resolution (VHR) are required to derive crop type and field extent [4,6,30]. Common sensors that acquire VHR data usually have the disadvantage of a limited number of spectral bands, which limits their capability when applying purely spectrally based classifications. Consequently, the incorporation of texture information contained in VHR data bears potential in data analysis [31–33]. In general, additional textural features result in the highest improvement of classification compared to other additional information sources like ancillary data, multitemporal datasets, or spectral indices [34].

Multispectral satellite data, e.g., stemming from RapidEye [35], SPOT 5 [36], QuickBird MS [16], IKONOS [16,37], WorldView-2 [38], or WorldView-3 [14], have been widely used to assess land cover in agricultural areas. These datasets have typical spatial resolutions between 0.5 m–6.5 m, and varying spectral bands (in terms of number, center wavelength and band width) mainly in the visible (VIS) and near-infrared (NIR) spectral range.

Nowadays, an increasing number of unmanned aerial vehicles (UAVs) provide the possibility to acquire VHR data. With such systems, it is possible to flexibly record data in an optimal spatial resolution. For agricultural mapping of small-scaled fields, a spatial resolution between 0.5 m–3 m is suggested to assess the infield variability [39]. Operating UAVs enables customized data acquisition in support of the definition and implementation of a dedicated processing chain for subsequent product generation.

Typical application fields of UAVs comprise vegetation, environmental, urban, or disaster monitoring [40], as well as precision agriculture, land cover mapping, and rangeland monitoring in the agricultural sector [41]. VHR data derived with a UAV have been used to determine land cover of small agricultural areas up to a few square kilometers [42]. Compared to satellite or airborne data acquisitions, UAVs are less expensive to operate [43,44], data acquisition needs less human interaction compared to destructive measurement methods [33], and is more flexible concerning weather conditions and flight planning [33,40]. Constraints are imposed by weather (e.g., rain and wind) and unfavorable solar illumination conditions. Nevertheless, since UAVs are usually operated at low flight altitude, data acquisition is also possible under cloudy conditions. In some countries, legal

restrictions for operating UAVs apply [43,45,46]. Apart from that, UAV data acquisition is more flexible throughout the day and not limited by given revisit times, as in the case of satellites, or potential flight restrictions with larger airborne platforms.

Although a wide range of sensors, ranging from consumer-grade cameras over multispectral sensors to imaging spectrometers [40], is available for UAVs today, their spectral calibration for sophisticated higher-level products requires additional effort [47,48]. On the one hand, sensors need to be calibrated while operated, e.g., by deploying a standard (white) reference in the field [49]. On the other hand, the exact spectral behavior of the sensor spectral bands must be known to eventually derive physical quantities [50].

In our study, we used a VHR dataset acquired with two uncalibrated consumer-grade cameras carried by a UAV for crop classification. While one camera captured the common red, green, and blue (RGB) bands, the other one recorded data with near-infrared, green, and blue (NirGB) bands. The data of the two cameras were combined into a NirRGB dataset. We present a novel method combining spectral and textural information to classify agricultural crops in a typical small structured arable landscape in the Swiss Plateau, using a random forest (RF) classifier [51] and VHR data from uncalibrated consumer-grade cameras. In this study, we analyze the influence of (i) spatial resolution, (ii) choice of spectral bands, and (iii) number of textural features, i.e., different sizes of the structuring element (SE), on the classification accuracy on a per-pixel basis and on the level of aggregated parcels.

2. Material

2.1. Study Area

The study area is situated in the Swiss Plateau within the Canton of Zurich (47.312°N, 8.733°E), Switzerland (Figure 1). The rural area is mainly covered by cropland and grassland. The elevation of the test site varies between 440 m–570 m above sea level (a.s.l.) and the climate can be described as warm temperate humid, with a yearly mean temperature around 9.3 °C and annual precipitation around 1134 mm [52]. Soils comprise mainly clay loam or loam, and Cambisol [53].

The predominant crop types in the study area are maize, sugar beet, and winter wheat (Table 1). Grassland comprises perennial (permanent) and annual (i.e., temporary) cover. Fields that were covered with hay during data acquisition and pure clover were treated as separate classes. The few and small spelt and winter barley fields were also taken into account. Rapeseed fields covered a minor area, and the bare soil fields were later planted with maize. Crop types present on less than three fields in the study area were excluded from our study. Additionally, the individual crop classes were grouped into generalized, merged classes to assess the performance of the subsequent classification (Table 1). Winter wheat, winter barley, and spelt were combined to cereals, the grassland class was merged with clover, and maize was extended to include pure soils, since maize was grown on these plots later during the year. Hay-covered fields were eventually excluded from our analysis, due to their heterogeneous appearance.

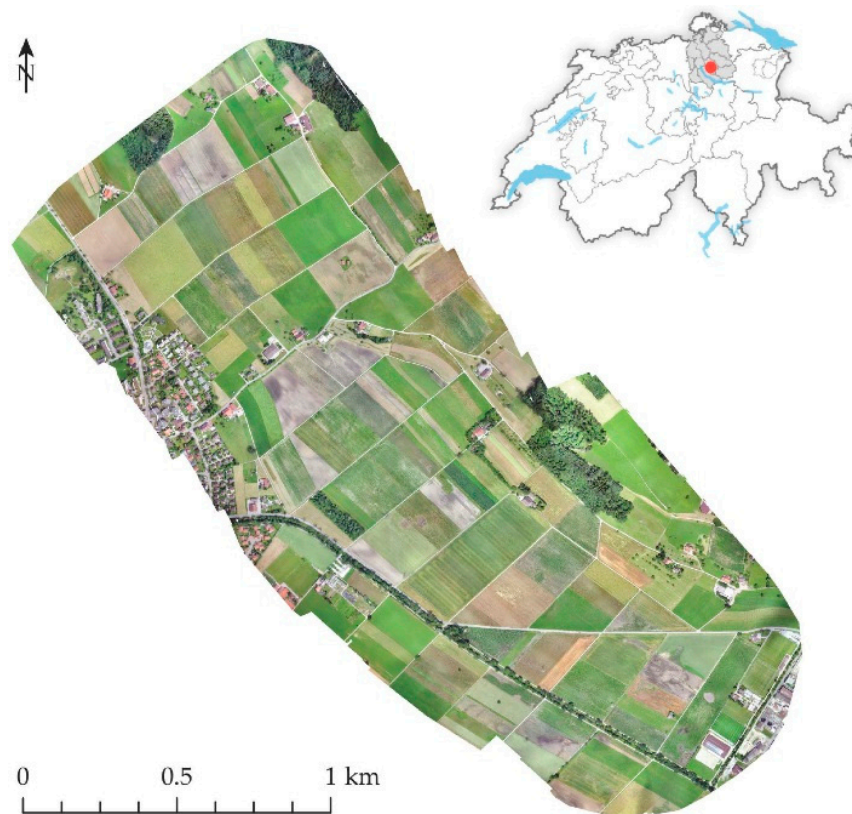


Figure 1. Outline of the study area (near Mönchaltorf, Switzerland). (See overview map of Switzerland, top right) in the Canton of Zurich, acquired with a spatial resolution of 0.05 m using a consumer-grade RGB (red, green, and blue) camera mounted on an unmanned aerial vehicle (UAV).

Table 1. Crop classes present in the study area and their respective field characteristics.

Crop Class		Total Area (ha)	Number of Fields	Spacing		Phenology (BBCH)
Full Set	Merged Set			Within-Row (cm)	Row (cm)	
Maize	Maize	19.6	15	14–16	75	0–33
Bare Soil		7.4	5	-	-	-
Sugar Beet	Sugar Beet	14.1	7	16	50	39
Winter Wheat	Cereals	24.4	13	5	14–15	75
Spelt		2.6	3	5	14–15	75
Winter Barley		2.5	3	5	14–15	99
Grassland	Grassland	15.0	17	-	-	-
Clover		5.4	3	-	10.5	-
Grass Hay	Excluded	3.6	5	-	-	-
Rapeseed	Rapeseed	7.6	6	10	30	80

2.2. Dataset

The entire study area comprises an extent of 170 ha, whereof 102 ha were taken into account. Data acquisition took place between 11:00–13:45 local time on 26 June 2015 in four flights of 30 min each. At this time of the year, all investigated crop types were present. The agricultural fields have an average size of 1.3 ha, ranging from 0.03 ha to 7.4 ha, with lengths of 140 m–200 m and widths of 23 m–180 m. Two larger fields have an extent of $183 \times 437 \text{ m}^2$ (8 ha) and $465 \times 87 \text{ m}^2$ (4 ha).

At the end of June 2015, the various crops were in different phenological stages. We determined the phenological code based on “Biologische Bundesanstalt, Bundessortenamt und CHemische Industrie”

(BBCH) [54] by non-destructive field inspections. Cereals were in maturity stage, with winter wheat and spelt in milk-ripe stage (BBCH 75), and winter barley in senescence (ready for harvest, BBCH 99). Maize included freshly sown to stem elongation stages (BBCH 0–33), rapeseed was just at the beginning of ripening (BBCH 80), and sugar beets had reached complete soil cover (BBCH 39). For grassland and clover, the exact phenological stages were not determined, since they were subject to a range of differing management. The phenological stages of pastures were heterogeneous due to grazing, whereas in the case of perennial and annual grasslands, phenological differences were linked to differing cutting strategies, reaching from complete mowing of the entire field to daily cuts of small parts for fresh forage.

Data acquisition was performed under clear sky conditions with a few condensation trails present. As the typical flight time of the deployed eBee UAV (Sensefly, Cheseaux-Lausanne, Switzerland) is limited to approximately 30 min, the total study area was divided into two parts in order for each subarea to be recorded in a single flight. For each subarea, both a 16.1 megapixel Canon IXUS 125HS camera with red, green, and blue (RGB) bands (center wavelengths at 660, 520, and 450 nm) and a modified camera of the same type with near infrared (NIR), green, and blue (NirGB) bands (center wavelengths at 720, 520, and 450 nm) were used consecutively. Flight planning and subsequent image acquisition were performed using the eMotion2 software (Sensefly, Cheseaux-Lausanne, Switzerland). The flight altitude was 150 m above ground, resulting in a spatial resolution of approximately 0.05 m. The images were acquired in parallel flight paths with a lateral overlap of 60% and a longitudinal overlap of 75%.

A total of 1092 single images were geo-tagged based on their respective GPS and IMU measurements on board the UAV during flight. The images were subsequently processed in Pix4D Mapper (Pix4D SA, Lausanne, Switzerland). The software uses the structure from motion (SfM) technique to generate a dense point cloud, a digital elevation model, and a mosaicked and rectified image product with a predefined spatial resolution of 0.05 m. During processing, five ground control points (GCP) that were measured with a differential GPS (dGPS) device on the ground were added for improved geo-rectification of the camera-wise image mosaics. The RGB mosaic and the NIR band was eventually stacked together with the RGB bands of the RGB camera, resulting in a VHR dataset consisting of four bands.

A crop type reference dataset of the study area was built based on a concurrent field survey and identified parcel boundaries (Figure 2). In order to avoid mixing effects at field borders, a buffer of 2 m was applied for classification training and validation (see Section 3).

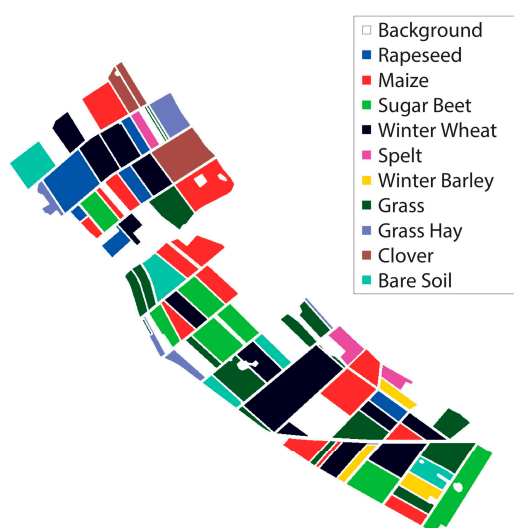


Figure 2. Reference map for all crop classes.

3. Method

We applied a robust classification and accuracy assessment workflow that consists of several steps (Figure 3). First, the VHR dataset was resampled to a range of spatial resolutions from which textural features were subsequently extracted. The features of these datasets were compiled to six different settings, based on spectral properties and amount of SE sizes. The data of these settings were then split into three parts for (i) training of the random forest (RF) model, (ii) validation of the model parameters, and (iii) testing of the final model classification performance. For the classification, we used an RF approach [51], having been widely used in previous studies and successfully applied [55]. The validated classification model was eventually applied to the test dataset, and an accuracy assessment was performed on both spatial supports (i.e., pixel- and parcel-based classification). The individual steps of our approach are described in detail below.

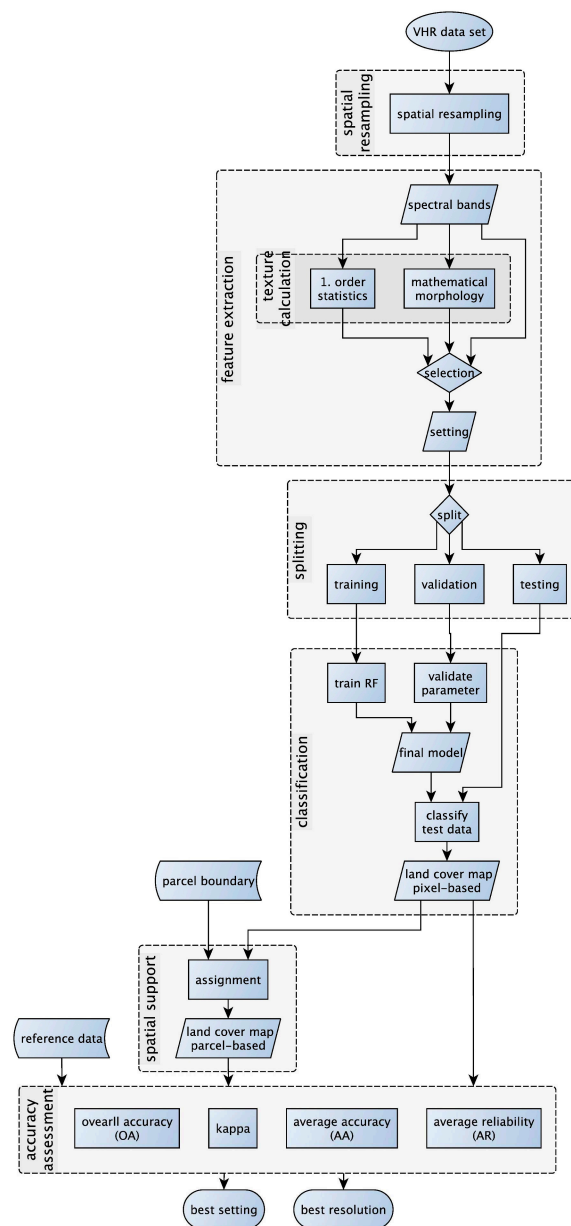


Figure 3. Flowchart of the method.

3.1. Resampling

To evaluate the influence of the spatial resolution on the classification accuracy, we resampled the VHR dataset to 0.1 m, 0.25 m, 0.5 m, 0.75 m, 1 m, and 2 m using a bicubic transformation. The reference dataset was resampled to the same spatial resolutions by applying a nearest neighbor method.

3.2. Feature Extraction

In order to incorporate spatial information into the classification chain, two types of textual features, i.e., first-order statistics and mathematical morphology, were calculated. The following statistical characteristics were used: mean, standard deviation, range, and entropy. Morphological operations comprised dilatation/erosion, opening/closing, opening/closing top hat, opening/closing by reconstruction, and opening/closing by reconstruction top hat [56–59]. The respective formulas can be found in Table 2.

These features were calculated based on a SE, i.e., a moving window. Since its shape and size are decisive, an SE is usually pre-selected based on expert knowledge. With some agricultural crops (in particular maize and sugar beet) in our dataset being cultivated in rows, their orientation has a major impact on the analyzed texture in an SE. Consequently, resulting feature values depend on the angle between plant rows and SE, especially in the case of a linear, but also a rectangular SE. To be rotation-invariant, all features were calculated in a disk-shaped SE.

The SE size, i.e., the diameter of the disk, was selected according to the measured distances between the plant rows in the study area, with the goal to include at least two rows of plants, in case of 5 cm spatial resolution. The distance between two rows for clover was 10.5 cm, for cereals 14 cm–15 cm, for rapeseed 30 cm, for sugar beet 50 cm, and for maize 75 cm (Table 1). To assess the texture of at least two crop rows, diameters of 3, 5, 9, 13, and 29 pixels were thus chosen as SE sizes. We applied the same SE sizes to all spatially resampled datasets.

For the subsequent analysis, six combinations of spectral bands and amount of SE sizes, so-called settings, were formed in total (Table 3). They are based on three sets of spectral bands, i.e., a set of all available bands of the two cameras (NIR, R, G, B), and two spectral subsets, representing the two cameras (NirGB and RGB) individually. Each of these spectral datasets was applied once to all SE sizes (5SE) and once to a reduced number of two SE sizes (2SE), with diameters of 3 and 5 pixels [38,60]. The added textural features were specifically built on the corresponding spectral subset. The settings are named according to the respective spectral bands and SE sizes (e.g., 5SE-NirRGB, comprising all spectral bands and all textural features with all SE sizes).

Table 2. Mathematical morphology formulae for image f and structuring element (SE) B for a pixel x . For further details see [56–59].

Name	Formula
Dilatation	$[\varepsilon Bf]_x = \min_{b \in B} f(x + b)$
Erosion	$[\delta Bf]_x = \max_{b \in B} f(x + b)$
Opening	$\Gamma Bf = \delta B \circ \varepsilon B(f)$
Closing	$\varphi Bf = \varepsilon B \circ \delta B(f)$
Opening by top hat	$OTH = f - \Gamma Bf$
Closing by top hat	$CTH = \varphi Bf - f$
Opening by reconstruction	$\gamma R(n) = Rf \delta [\varepsilon n f]$
Closing by reconstruction	$\varphi R(n) = Rf \varepsilon [\delta n(f)]$
Opening by reconstruction top hat	$ORTH = f - \gamma Rn(f)$
Closing by reconstruction top hat	$CRTH = \varphi Rn f - f$

Table 3. Composition of spectral and textural settings.

Setting	Spectral Bands	SE Sizes (Diameter (Pixel))
5SE-NirRGB	NIR-R-G-B	5SE (3, 5, 9, 13, 29)
5SE-RGB	R-G-B	5SE (3, 5, 9, 13, 29)
5SE-NirGB	NIR-G-B	5SE (3, 5, 9, 13, 29)
2SE-NirRGB	NIR-R-G-B	2SE (3, 5)
2SE-RGB	R-G-B	2SE (3, 5)
2SE-NirGB	NIR-G-B	2SE (3, 5)

3.3. Data Splitting for Validation

The dataset was split in three parts to perform a 3-fold cross validation, whereby one split was used for training of the RF classifier, one for validation of RF parameters, and the last one for testing to determine the classification accuracy. This ensures that only data that were not used for training and validation were used for classification [61]. Therefore, entire fields were assigned randomly to one of these data subsets, such that one third of the fields of a crop class was assigned to a split. In order to avoid these split specific assignments that influence the classification, all six possible permutations, called folds, were exercised.

3.4. Classification

To train the RF and to validate the model parameters, a set of 1000 stratified, randomly sampled pixels per class were selected from the respective training and validation datasets. The native TreeBagger implementation in MATLAB Version 2016a was used for the RF classifier. Usually, the number of trees is preselected by preliminary tests [38], or default values may be used [18]. In our case, we trained the RF with 20 logarithmically evenly spaced values between 10 and 1000 trees to determine the best amount of trees. A minimal leaf size of 3 was chosen to avoid overfitting. For all other parameters, default settings were kept, in particular, the square root of all features at each split.

In a first step (Equation (1)), we calculated the proportion of the correctly classified validation pixels and fitted them to an exponential function of the form

$$y = a * e^{-b*x} + c, \quad (1)$$

using a nonlinear least squares method with starting values of 0 for a and b, and 1 for c. A pre-study showed that this model and these parameters were the most suitable.

Then, we chose the number of trees with an accuracy loss of 0.1% compared to the best accuracy achieved with the fitted function in 1000 trees. However, in order to ensure stability, we set a threshold of at least 100 trees. Eventually, we trained the final model for classification with the determined number of trees, and all training and validation pixels. For the pixel-based classification, this final model was applied to all pixels of test data in the respective fold.

3.5. Spatial Support

Data smoothing at parcel level is commonly applied to agricultural classification results [16]. Parcels were resampled to the respective spatial resolutions using a nearest neighbor approach. The pixel-based classification was followed by the assignment of the most frequent label within a parcel to each pixel of the respective parcel, producing, thus, the parcel-based classification result.

3.6. Accuracy Assessment

The confusion matrix for the test dataset of each fold forms the basis for calculating the overall accuracy [62], kappa coefficient [63], as well as user and producer accuracy [62]. In order to get a better overall view, we averaged the values achieved for each fold. Overall accuracy (OA) refers to the average of the overall accuracy values of the six folds weighted by the number of total test pixels in

the corresponding fold. It is a very frequently used accuracy measure, and allows the comparison to other studies [34]. Kappa refers to the kappa coefficient, average accuracy (AA) refers to the mean of the user accuracies, and average reliability (AR) to the mean of the producer accuracies, respectively.

4. Results

We first present the overall best performing setting (i.e., 5SE-RGB) and then the influence of spatial resampling, selection of spectral bands, and amount of features, i.e., SE sizes. We show the results for the spatial support of pixels and parcels, as well as the full and merged sets of crop classes. All results over all classes are presented and discussed based on OA. Obtained kappa, AA, and AR are not described here in detail, although the corresponding values can be found in Tables S1–S4, and condensed summaries of accuracy values are given in Tables 4 and 5. Subsequently, we also present the class specific UA and PA values.

Table 4. Overall accuracy (OA (%)) for 5SE-RGB for all tested spatial supports (i.e., pixel- and parcel-based) and crop classes at different spatial resolutions.

Resolution (m)	Pixel-Based		Parcel-Based	
	Full Set	Merged Set	Full Set	Merged Set
0.1	61.1	80.2	76.1	94.7
0.25	60.5	86.5	69.9	96.7
0.5	66.7	86.3	74.0	94.6
0.75	63.7	86.5	65.7	96.2
1	62.6	86.0	61.9	94.3
2	60.0	82.7	67.8	92.2

Table 5. OA (%) for all settings, spatial supports and sets of crop classes at 0.5 m spatial resolution.

Settings	Pixel-Based		Parcel-Based	
	Full Set	Merged Set	Full Set	Merged Set
2SE-NirGB	53.5	72.4	68.0	79.4
2SE-NirRGB	60.5	77.1	75.6	83.5
2SE-RGB	60.1	81.5	65.0	92.6
5SE-NirGB	62.1	79.8	73.0	93.0
5SE-NirRGB	65.5	83.5	76.7	95.0
5SE-RGB	66.7	86.3	74.0	94.6

The best accuracy values are achieved for the 5SE-RGB setting at a spatial resolution of 0.5 m (Figure 4). In case of pixel-based classification the OA reaches 66.7%, whereas in the parcel-based case, an OA of 74.0% is achieved for the full set of crop classes (Table 4). The corresponding land cover maps can be found in Figure 5.

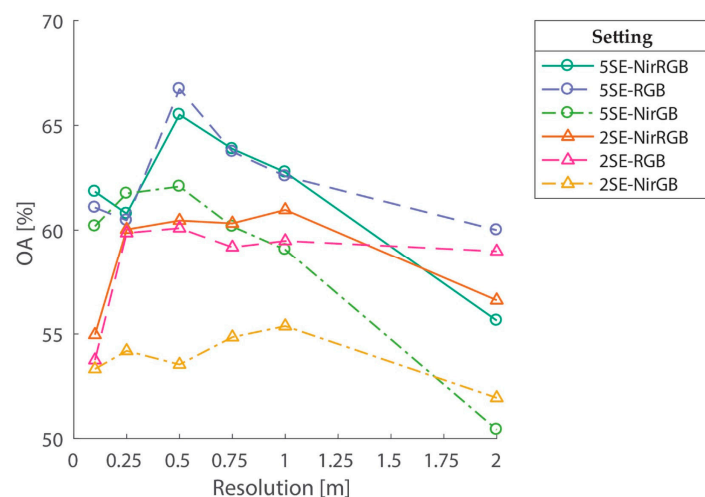


Figure 4. Overall accuracy (OA) for the full set of crop classes for all tested spatial resolutions and settings of the dataset for a pixel-based classification. Circles mark five structuring element (SE) sizes at evaluated resolutions and triangles two SEs, respectively. The line styles correspond to the applied spectral band selection.

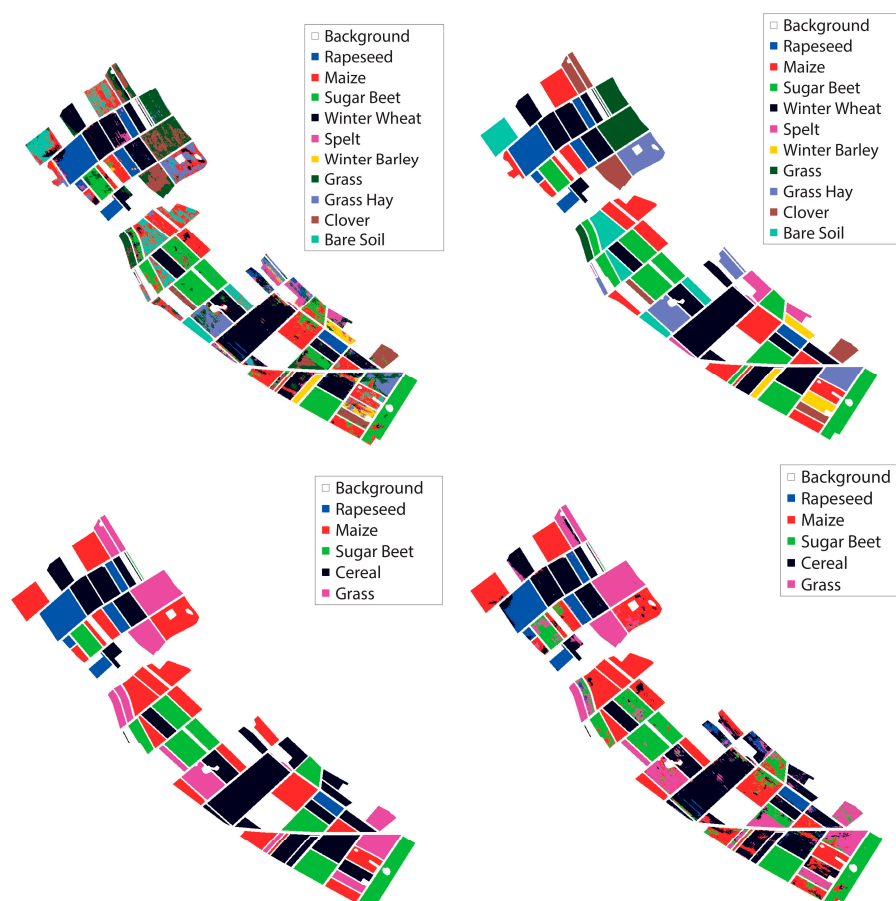


Figure 5. Land cover maps for the 5SE-RGB setting at 0.5 m spatial resolution. Full set of crop classes (**top**), merged set of crop classes (**bottom**), pixel-based classification (**left**), and parcel-based classification (**right**).

For the merged set of classes, the best values are achieved for 0.75 m spatial resolution for a pixel-based classification in terms of OA and AA (OA 86.5%, kappa 0.823, AA 85.8%, AR 87.9%), and for 0.25 m in terms of kappa and AR (OA 86.5%, kappa 0.823, AA 85.4.7%, AR 88.1%) (Table S3). However, the differences between the calculated accuracy measurements are small. The difference between the spatial resolution of 0.75 m and 0.5 m is 0.050% in terms of OA (differences in kappa: 0.000%, AA: 0.204%, AR: −0.127%). For the parcel-based classification, the best performance was achieved for a spatial resolution of 0.25 m with an OA of 96.7%, which is 2.1% better than for 0.5 m (Table S4).

4.1. Spatial Resampling

Concerning OA values for the tested spatial resolutions, most settings show a similar pattern for pixel-based classification of the full set of crop classes (Figure 4). The maximum OA is reached around 0.5 m, with decreasing values for higher or lower spatial resolution (Table 4).

In the case of pixel-based classification, the OA of the best performing setting (5SE-RGB) rises from 60.0% at a spatial resolution of 2 m, to 66.7% at 0.5 m for the full set of crop classes (Table 4). Then, it decreases to 61.1% at 0.1 m. For the set of merged classes, the OA raises from 82.7% at 2 m up to 86.3% at 0.5 m, and decreases to 80.2% at 0.1 m.

For parcel-based classifications, different spatial resolutions perform best for the full and merged set of crop classes (Figures S2 and S3). For the full set, a spatial resolution of 0.5 m yields the best classification accuracy, except for some settings that perform slightly better at 0.1 m spatial resolution. For the merged set of classes, highest accuracies are achieved at 0.25 m or 0.75 m spatial resolution, with slightly lower values at 0.5 m.

4.2. Spectral Resolution

Regarding spectral resolution of pixel-based classification, RGB settings generally lead to a better performance than settings with the additional NIR band (Figure 4 and Table 5). Settings without the red band (i.e., NirGB) perform worse. In the case of the full set of crop classes and the settings with five SE (5SE), this is only true for a spatial resolution of 0.5 m and 2 m (Table S1). For the other considered spatial resolutions, 5SE-NirRGB performs slightly better than 5SE-RGB in terms of OA. OA values for 5SE-NirGB are always lower than in the case of 5SE-NirRGB and 5SE-RGB, except for a spatial resolution of 0.25 m, where a 5SE-NirGB setting performs best. For settings with less SE sizes, 2SE-NirRGB outperforms 2SE-RGB, except for a resolution of 2 m. The OA values for classification of a 2SE-NirGB setting are always lower than for the other settings. In the case of the merged set of crop classes, the above statement applies to all spatial resolutions and amount of SE sizes (Figure S1).

Settings with RGB bands lead to better classification results than NirRGB and NirGB settings, irrespective of the spatial resolution and the amount of SE sizes (Table 5). On average, RGB settings show a 6% better OA compared to NIR-RGB, and 9% compared to NirGB for the full set of crop classes, and 7% and 10% in the case of the merged set, respectively.

In the parcel-based classification of the full set of crop classes, four-band settings perform best, followed by RGB and NirGB settings, with some exceptions for 0.1 m and 2 m spatial resolution (Table S2). Independent of the amount of SE sizes, 2SE-RGB and 5SE-RGB settings achieve better accuracies at a spatial resolution of 2 m than NirRGB settings. In addition, 5SE-RGB performs best at a spatial resolution of 0.1 m. In case of merged crop classes, RGB settings perform best, followed by NirRGB and NirGB settings.

4.3. Number of SE Sizes

In general, a higher number of SE sizes (5SE), and therefore, more textural features, lead to higher classification accuracies compared to a reduced amount (2SE) for the same spectral and spatial resolutions (Table 5). Nevertheless, 2SE settings outperform 5SE settings for parcel-based classification of the full set of crop classes at a spatial resolution of 0.75 and 1 m (Figure S1), and of the merged set of classes at 2 m (Figure S3), respectively.

4.4. Number of Classes and Spatial Support

In case of the best overall setting (i.e., 5SE-RGB, at 0.5 m), the classification result of the merged set of crop classes yields a 19.6% better OA compared to the full set for the pixel-based classification, and a 20.6% better OA for the parcel-based case (Table 5). The difference in OA between the two spatial supports is 7.3% for the full set of crop classes, and 8.3% for the merged set, respectively.

4.5. Class Specific Accuracy

For the full set of crop classes, the class specific accuracies range between 7.4% (UA of hay) and 100% (PA of various crops, e.g., sugar beet) for the best performing setting, i.e., 5SE-RGB at 0.5 spatial resolution, and a parcel-based classification (Tables S5 and S6). For the pixel-based classification, the range is slightly smaller and lies between 10.8% (UA of hay) and 91.8% (PA of rapeseed). Main mixtures occurred between maize and bare soil on the one hand and grassland, maize, and sugar beet on the other hand. The three cereal types mainly mixed up with each other. The same was true for grassland, clover and hay.

The class specific accuracies are slightly better with the additional NIR band in the NirRGB setting. Consequently, AA slightly increases to 60.0% and AR to 64.7% for the pixel-based classification (Table S1), and to 70.3% (AA) and 78.0% (AR) for the parcel-based classification (Table S1), compared to the 5SE-RGB setting. The mixtures between classes remain the same, but could be reduced. In particular, UA and PA of the cereals and bare soil could be improved by approximately 10%.

For the merged set of classes, the range of UA is 17.0% for the pixel-based classification and 14.8% for the parcel-based case, respectively. The range of PA is 15.5% for the pixel-based classification and 17.5% for the parcel-based case (Table S6). The primary mixtures occurred between grassland, maize, and sugar beet. In addition, rapeseed was mixed up with cereals, and the cereals with maize and grassland.

For some of the crop classes, a slight improvement of UA and/or PA could be achieved with the additional NIR band combined with a finer spatial resolution. For a spatial resolution of 0.25 m, AR increased by 1.5% for the parcel-based classification. The increase of AA with other settings and resolutions is negligible, as well as the increase of AA and AR for the pixel-based classification for the merged set of crops.

5. Discussion

A random forest-based classification method incorporating textural features was developed to assess the influence of spatial resolution, the choice of spectral bands, as well as the amount of different SE sizes on the classification accuracy of an uncalibrated, UAV-based VHR dataset. Overall, the best performing setting is 5SE-RGB at a spatial resolution of 0.5 m (Figure 4). For the full set of crop classes, an OA of 66.7% is achieved with a pixel-based classification. For a parcel-based classification, the OA increased by 7.3% to 74.0%. In the case of the merged set of crop classes, a similar behavior can be observed, with the OA for the pixel-based classification being 86.3%, and increasing by 8.3% to 94.6% for the parcel-based classification.

5.1. Influence of Spatial Resolution

Additional textural features, along with spectral data, improve the classification result, but these features depend on the spatial resolution of the sensor and SE size and number. This is consistent with [34], who found that additional textural features resulted in the highest improvement of a classification. In case of coarse resolutions, however, texture does not always improve the results, as was shown for mapping crops in an agricultural area in Austria with spatial structures similar to our study area and using Sentinel-2 data of 10 m resolution [18].

The spatial resolution of the dataset is crucial, because it determines the degree of detail. The elements that cause texture effects in crops are, on the one hand, the row spacing and the

within-row spacing of plants and, on the other hand, the visible bare soil in between the plants. At full canopy closure, the effects are mainly caused by shading of leaves and varying reflectance properties at different leaf angles [64]. In coarse resolution datasets (several meters), all these effects are integrated in the measurement of a single pixel, whereas in high-resolution data (few centimeters) the different leaf angles or even pebbles on the soil in the background are captured by a single pixel. Consequently, the optimal spatial resolution is driven by the fact that the between-class variability of pixels allows for the discrimination of the crop type, while not hampering the classification algorithm by within-class variation.

For industrially managed crops, the texture effects depend largely on the spacing between both the within-row and the row spacing [37]. Usually, single plants (or seeds) are placed at an optimal mutual distance in order to achieve the maximal possible yield [65], or rather, profit [66]. Therefore, best classification accuracies are achieved at an optimal spatial resolution where within-class variability of (texture) features is smaller than between-class variability [60].

In case of crops, both the within-class variability and the between-class variability decline with coarser spatial resolution. Consequently, the best performing spatial resolution is a trade-off in within- and between-class variability [67]. On the one hand, the within-class variability needs to be minimized. This is achieved when multiple plants are covered by a single pixel. On the other hand, the between-class variability should be as large as possible. For coarser resolutions, neighboring pixels in a class become more similar, and as a consequence, texture properties of different classes converge.

Within-row and row spacing influence the textural features of crops. In our study site, only sugar beet and maize are not yet in a stage of complete canopy closure at the end of June. Hence, the best spatial resolution is in the same range as the row spacing of these two classes. With a row spacing of 0.5 m for sugar beet, this value is equal to the best performing spatial resolution. Since maize fields are in (i) very early and (ii) heterogeneous phenological stages (Table 1), their row spacing does not have a dominant effect on the best spatial resolution.

Nevertheless, the optimal spatial resolution is also dependent on the spatial support, and number and kind of crop classes. For the pixel-based classification with a merged set of crop classes, the optimal spatial resolution is slightly coarser, i.e., a spatial resolution of 0.75 m performs best (Table 4). In contrast to the full set of crop classes, discrimination of grassland and clover, as well as among different cereals, is no longer performed. In case of parcel-based classification, a dataset based on 0.25 m spatial resolution yields higher classification accuracies than a dataset of 0.5 m pixel size (Figure S3).

In summary, a spatial resolution of 0.5 m performs best, in general, despite small accuracy losses for some classes or spatial supports. This is consistent with the findings of [39], where a spatial resolution of 0.5 m was found to be optimal to analyze the in-field variability of pasture using the red band of a multispectral sensor.

Numerous studies have been based on datasets of coarser spatial resolution acquired e.g., by the Moderate Resolution Imaging Spectrometer (MODIS) or Landsat, and providing sufficient spatial, spectral, and temporal resolution for large scale field monitoring [6]. Their main difference to our study relates to the prevalent field sizes in the Swiss Plateau, being smaller than elsewhere. In agricultural areas like, for instance, the US Central Great Plains, single fields are larger, with field sizes of more than 30 ha [13]. These areas are not as small-scaled as in Switzerland. Consequently, data of higher spatial resolution is necessary to analyze crop types in study areas like ours [4].

5.2. Impact of Spectral Characteristics

In general, remotely sensed data of spaceborne instruments are of more favorable spectral and radiometric specifications than VHR data obtained with an uncalibrated consumer-grade camera carried on a UAV, as in our study. Unlike our system that only acquires data in RGB bands and in an additional NIR band, datasets of e.g., Landsat 8 or MODIS provide a broader spectral range, with a number of bands in the NIR and SWIR spectral region. In addition, the spectral characteristics of spaceborne instruments are better defined in terms of spectral band width, and full width at

half maximum (FWHM). Nevertheless, our study demonstrates the feasibility to generate crop maps of documented accuracy, based on the respective VHR data and following the proposed classification method.

A number of studies document the benefits of using a NIR band [33,37] or a band in the red edge region [35] for crop classification. However, in our study, the NIR band does not improve the OA, in general. The NIR band of the modified Canon IXUS 125HS camera covers the wavelength region of approximately 690–730 nm, where the data values of vegetation and bare soil occur to be very similar in the acquired dataset. Further, the red band with its spectral range of approximately 640–680 nm is closely situated to the NIR band. Highest accuracies were thus achieved with an RGB band configuration. The fact that NirGB performs worst in general demonstrates the importance of the red band in our constellation. Indeed, we find that the differences in remotely sensed data values of vegetation and bare soil are most pronounced in this band. Consequently, RGB settings without a NIR band perform better overall. Only in the case of the merged set of crop classes on a parcel-based classification the 5SE-NirRGB setting achieves a slightly better OA than the 5SE-RGB setting. However, the classification of sugar beet and grassland would profit from the additional NIR band in terms of class specific accuracy (UA and PA, as mentioned in Section 4.5), but only in combination with a higher spatial resolution. Due to very similar spectral behavior of the three cereal crops (i.e., winter wheat, winter barley, and spelt) any additional spectral information may improve a classification performance.

5.3. Effect of Different SE sizes

Besides the spatial and spectral resolution having an influence on the classification result, more and larger SE sizes improve the classification accuracy. Morphological features keep or erase the elements in the SE that cause the texture by enlarging or erasing dark or bright elements [56]. As mentioned in Section 5.1, the main textural elements in crops are plants and bare soil. Depending on the sun position, shade causes dark parts. Since the different crops were tilled with different within-row and row spacing, the SE sizes must be defined in a way that they capture all present gaps [60]. Therefore, settings using five SE sizes (5SE) perform better than those taking only two SE sizes (2SE) into account, since the SE sizes should correspond to the present crops and their spacing.

5.4. Influence of Spatial Support

Object-based, i.e., parcel-based, classification improves the classification result [33], and is considered as the state-of-the-art in crop mapping [16]. In our case, it improves the pixel-based classification by 20% in terms of OA. The required field boundaries originate either from an additional data source [16], manual digitalization from scratch [38], or unsupervised segmentation [42]. For rural areas in the Swiss Plateau, a manual digitalization of individual parcels is feasible, since the field boundaries usually remain stable over several years.

5.5. Considerations about Acquisition Date and Temporal Resolution

An accurate classification of agricultural crops depends on a suitable point in time for data acquisition, since phenological stages of crops are changing rapidly [4]. By the end of June, all crops present in our study area were accrued, apart from maize. At this time of the year, most cultures are in their final stage of maturation, except for maize and sugar beet. Other studies also considered earlier and later acquisition dates, but concluded that maturity is the most promising phenological stage for a monotemporal analysis [16]. A later acquisition date (e.g., 30 August) leads to confusion, as some of the winter crops have already been harvested [18], while earlier dates may affect differentiation between bare soil and small plants [16,18]. In our dataset, this issue applies to maize being in an early phenological stage, and therefore, mixing up with bare soil. Additionally, the phenological variability among individual maize fields is large. In multitemporal analyses, datasets acquired before the end of July are reported to be the most important, with later datasets leading only to a minor improvement of

the classification result [15]. However, classification of a monotemporal dataset can achieve similar accuracies as in the case of a multitemporal dataset [18].

5.6. Comparison to Other Studies

When comparing our findings on the best performing spatial resolution of 0.5 m to other studies, not only spatial and spectral properties of the dataset and amount of different SE sizes for the textural features need to be considered. As could be seen in the differences in classification accuracy for the full and merged set of crop classes, the result depends as well on the actual classes, and the spatial support.

A recent study based on a multilevel classification in central Ontario, Canada, mainly aimed to differentiate tree species [42]. Maize, wheat, soybean, and alfalfa were classified as a side product in a parcel-based classification. The dataset was obtained using an eBee UAV as well, but with different cameras acquiring spectrally calibrated data. The study achieved an OA of 89% using a dataset acquired with the Parrot Sequoia sensor (green, red, NIR, and red edge bands, spatial resolution of 12.9 cm). Simultaneously, an RGB true color dataset was acquired (spatial resolution of 3.42 cm), serving as the basis to classify the crops with an OA of 83%. Finally, a Sony DSC-WX220 RGB camera was deployed (successor of the camera used in our study) to acquire a dataset with a spatial resolution of 3.52 cm. With this dataset, crop classification resulted in an OA of 81%. In addition to the spectral bands, the authors used texture and normalized difference vegetation index (NDVI) features, as well. With the method presented in our study and the 5SE-RGB setting, we achieve an OA of 94.6% at a spatial resolution of 0.5 m for the parcel-based classification of the merged classes. Based on a spatial resolution of 0.1 m, we still achieve an OA of 92.8%. Our approach performs slightly better, most likely due to the resampling to a coarser spatial resolution. Despite the additional calibration and NDVI feature, the OA accuracy of the aforementioned study is slightly lower.

Another classification study on simulated Sentinel-2 data from the Marchfeld region in Lower Austria achieved an OA of 76.5% [18]. The authors used an object-based method to classify seven agricultural cultures (carrots, maize, onions, soya, sugar beet, sunflower, and winter crops) based on spectral features only. The lower accuracy compared to our study was likely due to the unfavorable data acquisition date (30 August). At that point in time, winter crops were already harvested, and were therefore classified based on the spectral signature of bare soil and crop residuals. The high soil proportions in the harvested fields led to confusion with onion fields. Their pixel-based classification, however, performed better (OA of 83.2%), although with a higher variability in class specific accuracy, than in our case.

A further monotemporal study analyzed an Ikonos dataset of a rural area in Bursa, northwest Turkey, with a spatial resolution of 4 m acquired on 13 June [16]. Only R, G, B, and NIR bands of the dataset were used. An OA of 83.6% was achieved for a pixel-based classification of maize, pasture, rice, sugar beet, wheat, and tomato using an SVM-based method. A parcel-based classification leads to an improvement of 12.5% in OA. In comparison to our proposed method applied to the merged set of crop classes, this study achieved a slightly lower OA for the pixel-based classification, and a slightly better OA for the parcel-based classification approach.

Crop height is an additional parameter that can be derived from UAV data. In [68], the authors used the difference of the surface height between two acquisitions of RGB and NIR data with a spatial resolution of 0.8 m on 30 June and 21 October in Texas, USA. It could be shown that with crop height alone, the classification quality was limited, due to high variance even in single fields. Therefore, the authors used spectral, textural, and spatial features in addition, and reached an OA of 97.50% for an object-based method, and an OA of 78.52% for a pixel-based maximum likelihood (ML) classification. The OA accuracy was 2.5% lower when height information was not used. The reported land cover consisted of corn, cotton, sorghum, grass, bare soil, and wheat, being well comparable to our merged setting of crop types. The OA of our parcel-based classification is similar to the reported case without crop height information. In the pixel-based case, we achieved a roughly 10% better OA.

A further study was performed in July 2015 in the same area, using an RGB and a NIR camera to acquire data of five crop classes (i.e., cotton, corn, sorghum, soybean and watermelon) and five non-crop classes (i.e., impervious ground, bare soil, fallow, water, grass, and forest) at a spatial resolution of 0.35 m [69]. The authors tested a variety of groupings with pixel- and parcel-based classifications. The pixel-based classifications were based on a three-band setting (RGB) or a four-band setting (NirRGB). Based solely on spectral bands, the pixel-based classifications achieved OAs between 62% and 69% for the most comparable grouping containing all crops, and a single class for non-crops. In our setting, with textural features and the full set of crop classes, we achieved a comparable accuracy (OA of 66.7%). However, our merged set of crop classes outperformed all groupings of [69] with an OA of 86.3%. For the parcel-based classification, the authors additionally used vegetation indices (VIs), and statistical, geometrical, and textural features. In contrast to our study, they found an improvement of OA with an additional NIR band, which could be caused by the VIs that are based on the NIR band. They achieved OAs between 73% and 91%, depending on the setting and number of bands. Hence, their parcel-based results are slightly better when compared to our full set of crop classes (OA 74.0%), but slightly less accurate than our merged set (94.6%).

In another study at the same location, a NirRGB dataset of 0.4 m spatial resolution was upscaled to 1, 2, 4, 10, 15, and 30 m pixel sizes [70]. The authors classified cotton, sorghum, soybean, watermelon, non-crop vegetation, and non-vegetated area in the RGB dataset with an OA of 83.3%, and in the NirRGB dataset with an OA of 90.42% at a spatial resolution of 0.4 m. For coarser pixel sizes, the OA decreased to less than 70%. Compared to our best performing setting (5SE-RGB at 0.5 m spatial resolution) with an OA of 86.3%, they achieved similar accuracies in the RGB case. Again, with the additional NIR band and the implementation of VIs, their performance increased by 7%.

5.7. Limitations of Our Method

Both our method and the employed dataset have some limitations. Besides the spectral bands, the current approach relies mostly on textural features. Additional spectral or multitemporal features could improve the classification [34]. So far, there are only the spectral bands themselves incorporated, but further spectral features, such as spectral indices, could lead to performance improvements [31].

Since parcel-based classifications lead to higher accuracies, ancillary information about field borders is required. With this information not always being available, a conditional random field (CRF) smoothing to homogenize class assignments of a pixel-based classification for a certain field could alternatively be applied, leading to only slightly lower accuracies, but not being dependent on additional information sources [16]. Alternatively, a classification of segments within fields could improve the results as well [42].

A NIR band at ~800 nm instead of 720 nm would most likely improve the classification, as can be seen from the comparison of a Sequoia Parrot sensor and our modified Canon IXUS 125HS camera containing a NIR band [42]. This would allow an appropriate incorporation of vegetation indices (e.g., NDVI, generalized difference vegetation index (GDVI), or soil-adjusted vegetation index (SAVI)). Moreover, the spatial resolution of individual bands should be further analyzed, since a coarser resolution for NIR bands compared to a red band could lead to similar results [39].

6. Conclusions

We presented a classification method for crops on a dataset obtained by uncalibrated consumer-grade cameras mounted on a UAV. We analyzed different spatial and spectral resolutions, as well as different SE sizes for textural features. We investigated pixel and parcel-based spatial support and two sets of crop classes. On the one hand, we analyzed nine individual crop classes, and on the other hand, we pooled maize and bare soil, the three cereals types, and grassland and clover together.

Overall, the best performance was achieved with a dataset consisting of RGB bands and textural features of five structuring element (SE) sizes at a spatial resolution of 0.5 m. We were able to show that both a finer and a coarser spatial resolution perform worse. Settings that take the RGB bands into

account outperform such with the additional NIR band. Nevertheless, the NIR band leads partially to class specific improvements, but to slightly less accurate crop maps when all crops are classified together. SE sizes that cover the entire range of both within-row and row spacing of crops perform better. Consequently, our tested settings with five SE sizes outperform settings with two SE sizes.

A reduced set of crop classes led to better classification results (increase of ~7% in OA). As in other studies, we were not able to properly discriminate clover from grassland and the different cereal types from each other. Maize was in heterogeneous phenological stages ranging from fresh sown to stem elongation and could, therefore, not be distinguished from bare soil. As expected, parcel-based classification led to an improvement of ~20% in terms of OA compared to a pixel-based classification.

We conclude that a dataset with a spatial resolution of 0.5 m, consisting of spectrally poorly characterized and uncalibrated RGB bands, can provide sufficient information to differentiate between agricultural crop classes, given a set of SE sizes to describe textural features is taken into account in an appropriate manner. With an increasing availability of spaceborne VHR imagery becoming operationally available in the near future, the classification method presented and evaluated in this study contributes to the generation of crop maps of documented accuracy in small-scaled agricultural areas.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-4292/10/8/1282/s1>. Figure S1: Overall accuracy (OA) for the merged set of crop classes for all tested spatial resolutions and settings of the data set for a pixel-based classification. Circles mark five structuring element (SE) sizes at evaluated resolutions and triangles two SE, respectively. The line styles correspond to the applied spectral band selection. Figure S2: Overall accuracy (OA) for the full set of crop classes for all tested spatial resolutions and settings of the data set for a parcel-based classification. Circles mark five structuring element (SE) sizes at evaluated resolutions and triangles two SE, respectively. The line styles correspond to the applied spectral band selection. Figure S3: Overall accuracy (OA) for the merged set of crop classes for all tested spatial resolutions and settings of the data set for a parcel-based classification. Circles mark five structuring element (SE) sizes at evaluated resolutions and triangles two SE, respectively. The line styles correspond to the applied spectral band selection. Table S1: Accuracy values for all tested spatial resolutions and settings of the full set of crop classes for a pixel-based classification. Table S2: Accuracy values for all tested spatial resolutions and settings of the full set of crop classes for a parcel-based classification. Table S3: Accuracy values for all tested spatial resolutions and settings of the merged set of crop classes for a pixel-based classification. Table S4: Accuracy values for all tested spatial resolutions and settings of the merged set of crop classes for a parcel-based classification. Table S5: User Accuracy (UA) and Producer Accuracy (PA) for the full set of crop classes in a pixel- and parcel-based classification at a spatial resolution of 0.5 m and the 5SE-RGB setting. Table S6: User Accuracy (UA) and Producer Accuracy (PA) for the merged set of crop classes in a pixel- and parcel-based classification at a spatial resolution of 0.5 m and the 5SE-RGB setting.

Author Contributions: J.E.B. designed the research and analyzed the data with scientific advice of M.K. and M.E.S., J.E.B. wrote the manuscript and all co-authors thoroughly reviewed and edited the manuscript.

Funding: This research received no external funding.

Acknowledgments: The contribution of M.E.S. is supported by the University of Zurich Research Priority Program on ‘Global Change and Biodiversity’ (URPP GCB). We thank staff from RSL and GRS (Wageningen University) supporting fieldwork and providing background information.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rosenzweig, C.; Elliott, J.; Deryng, D.; Ruane, A.C.; Müller, C.; Arneth, A.; Boote, K.J.; Folberth, C.; Glotter, M.; Khabarov, N.; et al. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 3268–3273. [[CrossRef](#)] [[PubMed](#)]
2. Gerland, P.; Raftery, A.E.; Ševčíková, H.; Li, N.; Gu, D.; Spoorenberg, T.; Alkema, L.; Fosdick, B.K.; Chunn, J.; Lalic, N.; et al. World population stabilization unlikely this century. *Science* **2014**, *346*, 234–237. [[CrossRef](#)] [[PubMed](#)]
3. Tilman, D.; Balzer, C.; Hill, J.; Befort, B.L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20260–20264. [[CrossRef](#)] [[PubMed](#)]

4. Whitcraft, A.K.; Becker-Reshef, I.; Justice, C.O. A framework for defining spatially explicit earth observation requirements for a global agricultural monitoring initiative (GEOGLAM). *Remote Sens.* **2015**, *7*, 1461–1481. [[CrossRef](#)]
5. Fuhrer, J.; Gregory, P.J. *Climate Change Impact and Adaptation in Agricultural Systems*; CABI: Wallingford, CT, USA, 2014; ISBN 9781780642895.
6. Atzberger, C. Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sens.* **2013**, *5*, 949–981. [[CrossRef](#)]
7. Kastens, J.H.; Kastens, T.L.A.; Kastens, D.L.A.; Price, K.P.; Martinko, E.A.; Lee, R.Y. Image masking for crop yield forecasting using AVHRR NDVI time series imagery. *Remote Sens. Environ.* **2005**, *99*, 341–356. [[CrossRef](#)]
8. Zhang, P.; Anderson, B.; Tan, B.; Huang, D.; Myneni, R. Potential monitoring of crop production using a satellite-based Climate-Variability Impact Index. *Agric. For. Meteorol.* **2005**, *132*, 344–358. [[CrossRef](#)]
9. Mueller, N.D.; Gerber, J.S.; Johnston, M.; Ray, D.K.; Ramankutty, N.; Foley, J.A. Closing yield gaps through nutrient and water management. *Nature* **2012**, *490*, 254–257. [[CrossRef](#)] [[PubMed](#)]
10. Verbesselt, J.; Hyndman, R.; Newnham, G.; Culvenor, D. Detecting trend and seasonal changes in satellite image time series. *Remote Sens. Environ.* **2010**, *114*, 106–115. [[CrossRef](#)]
11. Rembold, F.; Atzberger, C.; Savin, I.; Rojas, O. Using low resolution satellite imagery for yield prediction and yield anomaly detection. *Remote Sens.* **2013**, *5*, 1704–1733. [[CrossRef](#)]
12. Galford, G.L.; Mustard, J.F.; Melillo, J.; Gendrin, A.; Cerri, C.C.; Cerri, C.E.P. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sens. Environ.* **2008**, *112*, 576–587. [[CrossRef](#)]
13. Wardlaw, B.D.; Egbert, S.L.; Kastens, J.H. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sens. Environ.* **2007**, *108*, 290–310. [[CrossRef](#)]
14. Palchowdhuri, Y.; Valcarce-Diñeiro, R.; King, P.; Sanabria-Soto, M. Classification of multi-temporal spectral indices for crop type mapping: A case study in Coalville, UK. *J. Agric. Sci.* **2018**, *156*, 24–36. [[CrossRef](#)]
15. Azar, R.; Villa, P.; Stroppiana, D.; Crema, A.; Boschetti, M.; Brivio, P.A. Assessing in-season crop classification performance using satellite data: A test case in Northern Italy. *Eur. J. Remote Sens.* **2016**, *49*, 361–380. [[CrossRef](#)]
16. Ozdarici-Ok, A.; Ok, A.; Schindler, K. Mapping of Agricultural Crops from Single High-Resolution Multispectral Images—Data-Driven Smoothing vs. Parcel-Based Smoothing. *Remote Sens.* **2015**, *7*, 5611–5638. [[CrossRef](#)]
17. Inglada, J.; Arias, M.; Tardy, B.; Hagolle, O.; Valero, S.; Morin, D.; Dedieu, G.; Sepulcre, G.; Bontemps, S.; Defourny, P.; et al. Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. *Remote Sens.* **2015**, *7*, 12356–12379. [[CrossRef](#)]
18. Immitzer, M.; Vuolo, F.; Atzberger, C. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens.* **2016**, *8*, 166. [[CrossRef](#)]
19. Shi, Y.; Thomasson, J.A.; Murray, S.C.; Pugh, N.A.; Rooney, W.L.; Shafian, S.; Rajan, N.; Rouze, G.; Morgan, C.L.S.; Neely, H.L.; et al. Unmanned Aerial Vehicles for High-Throughput Phenotyping and Agronomic Research. *PLoS ONE* **2016**, *11*, e0159781. [[CrossRef](#)] [[PubMed](#)]
20. Tattaris, M.; Reynolds, M.P.; Chapman, S.C. A Direct Comparison of Remote Sensing Approaches for High-Throughput Phenotyping in Plant Breeding. *Front. Plant Sci.* **2016**, *7*, 1131. [[CrossRef](#)] [[PubMed](#)]
21. Holman, F.H.; Riche, A.B.; Michalski, A.; Castle, M.; Wooster, M.J.; Hawkesford, M.J. High Throughput Field Phenotyping of Wheat Plant Height and Growth Rate in Field Plot Trials Using UAV Based Remote Sensing. *Remote Sens.* **2016**, *8*, 1031. [[CrossRef](#)]
22. Fuhrer, J.; Tendall, D.; Klein, T.; Lehmann, N.; Holzkämper, A. *Water Demand in Swiss Agriculture—Sustainable Adaptive Options for Land and Water Management to Mitigate Impacts of Climate Change*; Art-Schriftenreihe: Ettenhausen, Switzerland, 2013; Volume 59, Available online: <https://www.agroscope.admin.ch/agroscope/de/home/publikationen/suchen/reihen-bis-2013/art-schriftenreihe.html> (accessed on 14 August 2018).
23. De Wit, A.J.W.; Clevers, J.G.P.W. Efficiency and accuracy of per-field classification for operational crop mapping. *Int. J. Remote Sens.* **2004**, *25*, 4091–4112. [[CrossRef](#)]
24. Waldhoff, G.; Lussem, U.; Bareth, G. Multi-Data Approach for remote sensing-based regional crop rotation mapping: A case study for the Rur catchment, Germany. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *61*, 55–69. [[CrossRef](#)]

25. Long, T.B.; Blok, V.; Coninx, I. Barriers to the adoption and diffusion of technological innovations for climate-smart agriculture in Europe: Evidence from the Netherlands, France, Switzerland and Italy. *J. Clean. Prod.* **2016**, *112*, 9–21. [CrossRef]
26. Swiss Federal Office for Agriculture (FOAG). *Swiss Agricultural Policy*; FOAG: Berne, Switzerland, 2004; Available online: <https://www.cbd.int/financial/pes/swiss-pesagriculturalpolicy.pdf> (accessed on 14 August 2018).
27. Benton, T.G.; Vickery, J.A.; Wilson, J.D. Farmland biodiversity: Is habitat heterogeneity the key? *Trends Ecol. Evol.* **2003**, *18*, 182–188. [CrossRef]
28. Bundesrat, D.S. Verordnung über die Direktzahlungen an die Landwirtschaft. 2017, Volume 2013, pp. 1–152. Available online: <https://www.admin.ch/opc/de/classified-compilation/20130216/index.html> (accessed on 14 August 2018).
29. Akademien der Wissenschaften Schweiz Brennpunkt Klima Schweiz. *Grundlagen, Folgen, Perspektiven*; Swiss Academies Reports; 2016; Volume 11, pp. 1–128. Available online: <http://www.akademien-schweiz.ch> (accessed on 14 August 2018).
30. Vuolo, F.; Neugebauer, N.; Bolognesi, S.; Atzberger, C.; D’Urso, G. Estimation of Leaf Area Index Using DEIMOS-1 Data: Application and Transferability of a Semi-Empirical Relationship between two Agricultural Areas. *Remote Sens.* **2013**, *5*, 1274–1291. [CrossRef]
31. Laliberte, A.S.; Rango, A. Texture and Scale in Object-Based Analysis of Subdecimeter Resolution Unmanned Aerial Vehicle (UAV) Imagery. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 761–770. [CrossRef]
32. Marcos, D.; Volpi, M.; Kellenberger, B.; Tuia, D. Land cover mapping at very high resolution with rotation equivariant CNNs: Towards small yet accurate models. *ISPRS J. Photogramm. Remote Sens.* **2018**, in press. [CrossRef]
33. Whitehead, K.; Hugenholtz, C.H.; Myshak, S.; Brown, O.; LeClair, A.; Tamminga, A.; Barchyn, T.E.; Moorman, B.; Eaton, B. Remote sensing of the environment with small unmanned aircraft systems (UASs), Part 1: A review of progress and challenges. *J. Unmanned Veh. Syst.* **2014**, *2*, 69–85. [CrossRef]
34. Khatami, R.; Mountrakis, G.; Stehman, S.V. A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. *Remote Sens. Environ.* **2016**, *177*, 89–100. [CrossRef]
35. Zhang, H.; Li, Q.; Liu, J.; Shang, J.; Du, X.; McNairn, H.; Champagne, C.; Dong, T.; Liu, M. Image Classification Using RapidEye Data: Integration of Spectral and Textual Features in a Random Forest Classifier. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 5334–5349. [CrossRef]
36. Topaloglou, C.A.; Mylonas, S.K.; Stavrakoudis, D.G.; Mastorocostas, P.A.; Theocharis, J.B. Accurate crop classification using hierarchical genetic fuzzy rule-based systems. In Proceedings of the SPIE—Remote Sensing for Agriculture, Ecosystems, and Hydrology, Amsterdam, The Netherlands, 22–23 September 2014; Volume 9239.
37. Helmholtz, P.; Rottensteiner, F.; Heipke, C. Semi-automatic verification of cropland and grassland using very high resolution mono-temporal satellite images. *ISPRS J. Photogramm. Remote Sens.* **2014**, *97*, 204–218. [CrossRef]
38. Debats, S.R.; Luo, D.; Estes, L.D.; Fuchs, T.J.; Caylor, K.K. A generalized computer vision approach to mapping crop fields in heterogeneous agricultural landscapes. *Remote Sens. Environ.* **2016**, *179*, 210–221. [CrossRef]
39. Atkinson, P.M. Selecting the spatial resolution of airborne MSS imagery for small-scale agricultural mapping. *Int. J. Remote Sens.* **1997**, *18*, 1903–1917. [CrossRef]
40. Pajares, G. Overview and Current Status of Remote Sensing Applications Based on Unmanned Aerial Vehicles (UAVs). *Photogramm. Eng. Remote Sens.* **2015**, *81*, 281–330. [CrossRef]
41. Shahbazi, M.; Théau, J.; Ménard, P. Recent applications of unmanned aerial imagery in natural resource management. *GISci. Remote Sens.* **2014**, *51*, 339–365. [CrossRef]
42. Ahmed, O.S.; Shemrock, A.; Chabot, D.; Dillon, C.; Williams, G.; Wasson, R.; Franklin, S.E. Hierarchical land cover and vegetation classification using multispectral data acquired from an unmanned aerial vehicle. *Int. J. Remote Sens.* **2017**, *38*, 2037–2052. [CrossRef]
43. Colomina, I.; Molina, P. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. [CrossRef]

44. Hugenholtz, C.H.; Moorman, B.J.; Riddell, K.; Whitehead, K. Small unmanned aircraft systems for remote sensing and Earth science research. *Eos Trans. Am. Geophys. Union* **2012**, *93*, 236. [CrossRef]
45. Whitehead, K.; Hugenholtz, C.H.; Myshak, S.; Brown, O.; LeClair, A.; Tamminga, A.; Barchyn, T.E.; Moorman, B.; Eaton, B. Remote sensing of the environment with small unmanned aircraft systems (UASs), Part 2: Scientific and commercial applications 1. *J. Unmanned Veh. Syst.* **2014**, *2*, 86–102. [CrossRef]
46. Stöcker, C.; Bennett, R.; Nex, F.; Gerke, M.; Zevenbergen, J. Review of the Current State of UAV Regulations. *Remote Sens.* **2017**, *9*, 459. [CrossRef]
47. Kelcey, J.; Lucieer, A. Sensor correction and radiometric calibration of a 6-band multispectral imaging sensor for UAV remote sensing. In Proceedings of the 12th Congress of the International Society for Photogrammetry and Remote Sensing, Melbourne, Australia, 25 August–1 September 2012; Volume 39, pp. 393–398.
48. Lu, B.; He, Y. Species classification using Unmanned Aerial Vehicle (UAV)-acquired high spatial resolution imagery in a heterogeneous grassland. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 73–85. [CrossRef]
49. Hueni, A.; Damm, A.; Kneubuehler, M.; Schläpfer, D.; Schaepman, M.E. Field and Airborne Spectroscopy Cross Validation—Some Considerations. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 1117–1135. [CrossRef]
50. Del Pozo, S.; Rodríguez-Gonzálvez, P.; Hernández-López, D.; Felipe-García, B. Vicarious Radiometric Calibration of a Multispectral Camera on Board an Unmanned Aerial System. *Remote Sens.* **2014**, *6*, 1918–1937. [CrossRef]
51. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
52. MeteoSwiss Climate Normals Zürich/Fluntern Climate Normals Zürich/Fluntern. Available online: http://www.meteoswiss.admin.ch/product/output/climate-data/climate-diagrams-normal-values-station-processing/SMA/climsheet_SMA_np8110_e.pdf (accessed on 19 March 2018).
53. Diek, S.; Schaepman, M.; de Jong, R. Creating Multi-Temporal Composites of Airborne Imaging Spectroscopy Data in Support of Digital Soil Mapping. *Remote Sens.* **2016**, *8*, 906. [CrossRef]
54. Meier, U. Entwicklungsstadien mono-und dikotylar Pflanzen. *Die erweiterte BBCH Monogr.* **2001**, *2*. Available online: http://www.agrometeo.ch/sites/default/files/u10/bbch-skala_deutsch.pdf (accessed on 14 August 2018).
55. Fernández-Delgado, M.; Cernadas, E.; Barro, S.; Amorim, D. Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? *J. Mach. Learn. Res.* **2014**, *15*, 3133–3181.
56. Fauvel, M.; Tarabalka, Y.; Benediktsson, J.A.; Chanussot, J.; Tilton, J.C. Advances in Spectral-Spatial Classification of Hyperspectral Images. *Proc. IEEE* **2013**, *101*, 652–675. [CrossRef]
57. Tuia, D.; Pacifici, F.; Kanevski, M.; Emery, W.J. Classification of Very High Spatial Resolution Imagery Using Mathematical Morphology and Support Vector Machines. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3866–3879. [CrossRef]
58. Vincent, L. Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms. *IEEE Trans. Image Process.* **1993**, *2*, 176–201. [CrossRef] [PubMed]
59. Soille, P. *Morphological Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2004; ISBN 978-3-642-07696-1.
60. Ferro, C.J.S.; Warner, T.A. Scale and texture in digital image classification. *Photogramm. Eng. Remote Sens.* **2002**, *68*, 51–63.
61. Lyons, M.B.; Keith, D.A.; Phinn, S.R.; Mason, T.J.; Elith, J. A comparison of resampling methods for remote sensing classification and accuracy assessment. *Remote Sens. Environ.* **2018**, *208*, 145–153. [CrossRef]
62. Congalton, R. *Assessing the Accuracy of Remotely Sensed Data*; Mapping Science; CRC Press: Boca Raton, FL, USA, 2008; ISBN 978-1-4200-5512-2.
63. Cohen, J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychol. Bull.* **1968**, *70*, 213–220. [CrossRef] [PubMed]
64. Norman, J.; Welles, J. Radiative transfer in an array of canopies. *Agron. J.* **1983**, *75*, 481–488. [CrossRef]
65. Flénet, F.; Kiri, J.R.; Board, J.E.; Westgate, M.E.; Reicosky, D.A. Row spacing effects on light extinction coefficients of corn, sorghum, soybean, and sunflower. *Agron. J.* **1996**, *88*, 185–190. [CrossRef]
66. Seginer, I.; Ioslovich, I. Optimal spacing and cultivation intensity for an industrialized crop production system. *Agric. Syst.* **1999**, *62*, 143–157. [CrossRef]
67. Tokarczyk, P.; Wegner, J.D.; Walk, S.; Schindler, K. Features, Color Spaces, and Boosting: New Insights on Semantic Classification of Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 280–295. [CrossRef]

68. Wu, M.; Yang, C.; Song, X.; Hoffmann, W.C.; Huang, W.; Niu, Z.; Wang, C.; Li, W. Evaluation of orthomosaics and digital surface models derived from aerial imagery for crop type mapping. *Remote Sens.* **2017**, *9*, 239. [[CrossRef](#)]
69. Zhang, J.; Yang, C.; Song, H.; Hoffmann, W.C.; Zhang, D.; Zhang, G. Evaluation of an Airborne Remote Sensing Platform Consisting of Two Consumer-Grade Cameras for Crop Identification. *Remote Sens.* **2016**, *8*, 257. [[CrossRef](#)]
70. Zhang, J.; Yang, C.; Zhao, B.; Song, H.; Hoffmann, W.C.; Shi, Y.; Zhang, D.; Zhang, G. Crop classification and LAI estimation using original and resolution-reduced images from two consumer-grade cameras. *Remote Sens.* **2017**, *9*, 2054. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).